

cn. us 6,094,636

AUDIO DATA ENCODING AND DECODING METHOD AND DEVICE WITH ADJUSTABLE BIT RATE

Patent number: JP10285042

Publication date: 1998-10-23

Inventor: KIM YEON-BAE

Applicant: SAM SUNG ELECTRONIC

Classification:

- International: G10L19/00; G10L19/02; G10L19/14; H04B1/66;
H04N7/26; H04N7/30; G10L19/00; H04B1/66;
H04N7/26; H04N7/30; (IPC1-7): H03M7/30; G10L9/00

- european: G10L19/00E; G10L19/02Q2; G10L19/14A1R;
H04B1/66M; H04N7/26A8T; H04N7/30

Application number: JP19980048443 19980227

Priority number(s): KR19970012233 19970402; KR19970061300 19971119

Also published as:



EP0869622 (A2)



US6094636 (A1)



KR261254 (B1)



EP0869622 (A3)



BR9705838 (A)

[more >>](#)

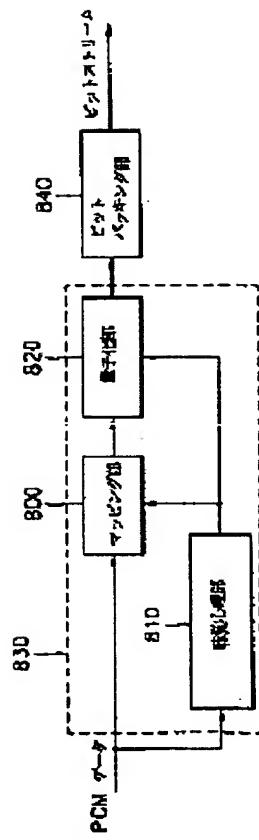
[Report a data error here](#)

Abstract of JP10285042

PROBLEM TO BE SOLVED: To adjust bit rate by encoding the quantization data belonging to an improvement hierarchy following an encoded hierarchy and also the quantization data which belong to the encoded hierarchy but are still remaining in each prescribed hierarchy size when the audio signals are encoded into a data stream.

SOLUTION: A mapping part 800 of a quantization processing part 830 converts the input audio signals of a time area into the signals of a frequency area, and an aural psychology part 810 bundles the converted signals into the prescribed area signals and calculates the masking threshold in every band. A quantization part 820 quantizes the audio signals in every prescribed encoding band, and a bit packing part 840 generates a bit stream from the data value on the quantized frequency band in response to the bit rate corresponding to every hierarchy.

Then the quantized data belonging to the improvement hierarchy following an encoded hierarchy and also the quantized data which belong to the encoded hierarchy but are still remaining are encoded for all hierarchies in each prescribed hierarchy size.



Data supplied from the esp@cenet database - Worldwide

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-285042

(43)公開日 平成10年(1998)10月23日

(51)Int.Cl.⁶

H 03 M 7/30
G 10 L 9/00

識別記号

F I

H 03 M 7/30
G 10 L 9/00

A
J

審査請求 未請求 請求項の数21 OL (全 16 頁)

(21)出願番号 特願平10-48443

(22)出願日 平成10年(1998)2月27日

(31)優先権主張番号 1997 12233

(32)優先日 1997年4月2日

(33)優先権主張国 韓国 (KR)

(31)優先権主張番号 1997 61300

(32)優先日 1997年11月19日

(33)優先権主張国 韓国 (KR)

(71)出願人 390019839

三星電子株式会社

大韓民国京畿道水原市八達区梅林洞416

(72)発明者 金 延培

大韓民国京畿道水原市權善區權善洞(番地
なし)新東亞アパート504棟306號

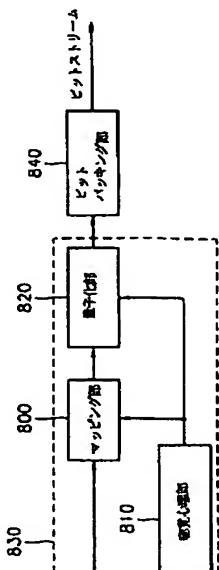
(74)代理人 弁理士 志賀 正武 (外1名)

(54)【発明の名称】 ピット率の調節可能なオーディオデータ符号化/復号化方法及び装置

(57)【要約】

【課題】 ピット率の調節可能なオーディオ符号化/復号化方法及び装置を提供することを目的とする。

【解決手段】 オーディオ信号を基本階層と所定数の向上階層よりなる階層的なデータストリームに符号化する方法は、入力オーディオ信号を信号処理して所定の符号化帯域別に量子化する量子化処理段階と、基本階層に応する量子化されたデータを所定の階層サイズ内で符号化する基本階層符号化段階と、符号化された階層の次の向上階層に属する量子化されたデータ及び前記符号化された階層に属しながら符号化されなく残っている量子化されたデータを所定の階層サイズ内で符号化する階層符号化段階と、階層符号化段階を全ての階層に対して行う順次符号化段階とを備える。



【特許請求の範囲】

【請求項 1】 オーディオ信号を基本階層と所定数の向上階層よりなる階層的なデータストリームに符号化する方法において、

- (a) 入力オーディオ信号を信号処理して所定の符号化帯域別に量子化する量子化処理段階と、
- (b) 前記基本階層に相応する量子化されたデータを所定の階層サイズ内で符号化する基本階層符号化段階と、
- (c) 前記符号化された階層の次の向上階層に属する量子化されたデータ及び前記符号化された階層に属しながら符号化されなく残っている量子化されたデータを所定の階層サイズ内で符号化する階層符号化段階と、
- (d) 前記階層符号化段階を全ての階層に対して行う順次符号化段階とを含むことを特徴とし、
前記(b)、(c)及び(d)段階の符号化は、
- (e) 符号化しようとする該当階層に属する各帯域に対して、前記帯域に相応する量子化されたデータのビット数を示す全域ビット割当情報を求める段階と、
- (f) 前記階層の各帯域の大きさの内で各帯域別に割当られるビット数を求める段階と、
- (g) 前記割当されたビット数に相応する量子化されたデータに対して前記帯域を構成している所定周波数成分別に量子化されたデータの存在有無を示す位置情報を生成する段階と、
- (h) 前記全域ビット割当情報、量子化ステップサイズ、位置情報及び前記帯域別に割当されたビット数に相応する量子化されたデータを所定の符号化方法によりビットストリームに生成する段階よりなるビット率の調節可能なオーディオデータ符号化方法。

【請求項 2】 前記帯域別に割当されたビット数に相応する量子化されたデータは、
量子化されたデータの最上位ビットから下位ビット順に相応する量子化されることであることを特徴とする請求項 1 に記載のビット率の調節可能なオーディオデータ符号化方法。

【請求項 3】 前記(h)段階の前記帯域別に割当されたビット数に相応する量子化されたデータの符号化は、前記量子化されたデータを所定数のビット単位に束ねて符号化することを特徴とする請求項 2 に記載のビット率の調節可能なオーディオデータ符号化方法。

【請求項 4】 前記(f)段階の割当されたビット数は1ビットであり、
前記(h)段階は、
前記全域ビット割当情報と位置情報を所定の符号化方法によりビットストリームに生成することを特徴とする請求項 2 に記載のビット率の調節可能なオーディオデータ符号化方法。

【請求項 6】 前記無損失符号化はハフマン符号化であることを特徴とする請求項 5 に記載のビット率の調節可能なオーディオデータ符号化方法。

【請求項 7】 前記無損失符号化は算術符号化であることを特徴とする請求項 5 に記載のビット率の調節可能なオーディオデータ符号化方法。

【請求項 8】 前記(a)段階は、
時間領域の入力オーディオ信号を周波数領域の信号に変換する段階と、

前記時間/周波数マッピングにより周波数成分に変換された入力オーディオ信号を所定の帯域信号に束ね、前記各帯域でのマスキングスレショルドを計算する段階と、各帯域の量子化ノイズがマスキングスレショルドより小さくなるように所定の符号化帯域別に量子化する段階よりなることを特徴とする請求項 1 に記載のビット率の調節可能なオーディオデータ符号化方法。

【請求項 9】 オーディオ信号を基本階層と所定数の向上階層よりなるビット率よりなる階層的なビット率を有するように符号化する装置において、

入力オーディオ信号を信号処理して所定の符号化帯域別に量子化する量子化処理部と、

前記基本階層に属する各帯域の量子化されたデータのビット数を示す基本階層の帯域別全域ビット割当情報及び前記各帯域別に割当られるビット数に相応する量子化されたデータの所定の周波数成別データの存在有無を示す位置情報、量子化ステップサイズ及び量子化されたデータを符号化し、前記基本階層に対する符号化が終わると、その次の階層に対する全域ビット割当情報、位置情報、量子化ステップサイズ及び量子化されたデータを符号化し、これを全ての階層に対して行ってビットストリームを形成するビットパッキング部とを含むことを特徴とするビット率の調節可能なオーディオデータ符号化装置。

【請求項 10】 前記帯域別に割当されたビット数に相応する量子化されたデータの符号化は、
最上位ビットから下位ビット順に行うことを特徴とする請求項 9 に記載のビット率の調節可能なオーディオデータ符号化装置。

【請求項 11】 前記ビットパッキング部の帯域別に割当されたビット数に相応する量子化されたデータの符号化は、
前記量子化されたデータを所定数のビット単位に束ねて符号化することを特徴とする請求項 10 に記載のビット率の調節可能なオーディオデータ符号化装置。

【請求項 12】 前記ビットパッキング部の帯域別に割当されたビット数は1ビットであり、
前記符号化は全域ビット割当情報と位置情報を所定の符

【請求項13】 前記ビットパッキング部の符号化は低周波数成分から高周波数成分順に行なうことを特徴とする請求項10に記載のビット率の調節可能なオーディオデータ符号化装置。

【請求項14】 前記量子化処理部は、時間領域の入力オーディオ信号を周波数領域の信号に変換する時間/周波数マッピング部と、前記時間/周波数マッピングにより周波数成分に変換された入力オーディオ信号を所定の帯域信号に束ね、前記各帯域でのマスキングスレショルドを計算する聴覚心理部と、各帯域の量子化ノイズがマスキングスレショルドより小さくなるように所定の符号化帯域別に量子化する量子化部とを含むことを特徴とする請求項9または請求項10に記載のビット率の調節可能なオーディオデータ符号化装置。

【請求項15】 階層的なビット率を有するように符号化されたオーディオデータを復号化する方法において、前記階層に属する各帯域に相応する量子化されたデータのビット数を示す全域ビット割当情報、量子化ステップサイズ、前記割当されたビット数に相応する量子化されたデータの周波数成分別存在有無を示す位置情報及び量子化されたデータを前記ビットストリームの階層が生成された順序通り復号化するが、前記各階層に属する帯域に割当されたビット数を求めて前記割当されたビット数に相応する量子化されたデータを前記位置情報に参照して復号化する復号化段階と、

前記復号化された量子化ステップサイズと量子化されたデータを元の大きさの信号に復元する段階と、前記逆量子化された周波数領域のオーディオ信号を時間領域の信号に変換する段階とを含むことを特徴とするビット率の調節可能なオーディオデータの復号化方法。

【請求項16】 前記復号化段階の量子化されたデータの復号化は最上位ビットから下位ビット順に行われるこことを特徴とする請求項15に記載のビット率の調節可能なオーディオデータの復号化方法。

【請求項17】 前記復号化段階の量子化されたデータの復号化は所定数のビットよりなるペクトル単位で復号化することを特徴とする請求項15に記載のビット率の調節可能なオーディオデータの復号化方法。

【請求項18】 前記復号化段階における復号化は算術復号化方法により復号化されることを特徴とする請求項15に記載のビット率の調節可能なオーディオデータの復号化方法。

【請求項19】 前記復号化段階における復号化はハフマン復号化方法により復号化されることを特徴とする請求項15に記載のビット率の調節可能なオーディオデータ

前記階層に属する各帯域に相応する量子化されたデータのビット数を示す全域ビット割当情報、量子化ステップサイズ、前記割当されたビット数に相応する量子化されたデータの周波数成分別存在有無を示す位置情報及び量子化されたデータを前記ビットストリームの階層が生成された順序通り復号化するが、前記各階層に属する帯域に割当されたビット数を求めて前記割当されたビット数に相応する量子化されたデータを前記位置情報に参照して復号化するビットストリーム分析部と、前記ビットストリーム分析部から復号化された量子化ステップサイズと量子化されたデータを元の大きさの信号に復元する逆量子化部と、前記逆量子化部から逆量子化された周波数領域のオーディオ信号を時間領域の信号に変換する周波数/時間マッピング部とを含むことを特徴とするビット率の調節可能なオーディオデータの復号化装置。

【請求項21】 ビットストリーム分析部の前記復号化段階の量子化されたデータの復号化は量子化されたデータの最上位ビットから下位ビット順に符号化されたデータを復号化することを特徴とする請求項20に記載のビット率の調節可能なオーディオデータの復号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明はデータ符号化及び復号化に係り、特にビット率の調節可能なデジタルオーディオデータの符号化/復号化装置及びその方法に関する。特に、一つのビットストリーム内に1個のビット率に対するのみで構成することでなく、基本段階(Base Layer)に基づき多数の上位段階(Enhancement Layer)のビット率に対したデータを共に表現するオーディオ符号化/復号化装置及び方法に関する。

【0002】

【従来の技術】 オーディオ機器は記録貯蔵媒体に信号を貯蔵した後、使用者が必要時に貯蔵された信号を聞く装置である。最近、デジタル信号処理技術の発達により既存のアナログ信号によるLP及びテープからデジタル信号によるCD(Compact Disc)とDAT(Digital Audio Tape)への開発が進行されて音質の向上を遂げたが、データの量が多くて貯蔵及び伝送に問題があった。従ってデータの量を減らすためにDPCM(Differential Pulse Code Modulation)やADPCM(Adaptive Differential Pulse Code Modulation)等の方法が提案されたが、信号の種類に応じて効率性が大きく異なる短所があった。最近、ISO(International Standard Organization)により標準化作業がなされたMPEG(Moving Pictures Expert Group)オーディオや商品化されたフィリップスのDCC(Digital Compact Cassette)、SONYのMD(Mini Disc)では人間の聴覚心理(psy

【0003】人の聴覚心理特性を考慮したオーディオ符号化装置の構成は図1のように時間/周波数マッピング部100、聴覚心理部110、ビット割当部120、割当てられたビットにより量子化処理する量子化(quantization)部130、ビットパッキング部(bitpacking)140よりなる。ここで、前記聴覚心理部110は人間の聴覚特性を用いて信号対マスキング比を計算するブロックである。即ち、人間の聴覚特性のうち特にマスキング現象を用いたことであって、入力信号の大きさと信号の相互作用により人間が聞いても感じられない信号の最小値のマスクされたスレショルド(Masked Threshold)を求める。ビット割当部120はマスクされたスレショルドを用いて聴覚に重要な役割をする信号がある部分から限定されたビット内で先にビットを割当てられることによりデータの圧縮効果を得る部分である。

【0004】人間の聴覚特性のうちデジタルオーディオ信号の符号化において最も重要な現象として使用するのがマスキング(masking)効果と臨界帯域(critical band)の性質である。マスキング効果とは何れの信号(音)により他の信号(音)が聞こえない現象である。図2は前記マスキング現象を説明するための図面であって、例えば駅で列車が通る時小声で話合うと、その声は列車音に葬られて聞こえなくなる現象である。臨界帯域は人間の可聴周波数範囲内で何れの帯域別に信号を認知する現象であるが、様々な特性を有している。これら特徴の中の1つは同じ大きさのノイズに対して1つの臨界帯域内にある時と臨界帯域範囲を越える時、人間が認知するノイズの大きさが異なるということである。この場合、臨界帯域を外れる時のノイズに対する認知がさらに高まる。

【0005】聴覚特性を用いて符号化することは基本的にこの2つの特性を用いて1つの臨界帯域内でどのくらいノイズが割当てられるかを計算した後、量子化時量子化ノイズをそれだけ発生させることにより符号化による情報の損失を最小化することである。このようなデジタルオーディオ符号化方法の応用はデジタルオーディオ放送やオーディオ機器、マルチメディア、AOD(Audio On Demand)などがある。このような符号化方法はほとんど固定ビット率(fixed bitrate)を支援する。即ち、一つのビット文字列(bitstream)が1個の特定ビット率(例えば、128kbps、96kbpsまたは64kbps)に対して構成されている。このような構成は伝送線がオーディオデータ専用線として使われる時は何ら問題がない。専用線とは特定ビット率を固定的に支援する線路であるため、専用線に適合する特定ビット率で構成されたビット文字列は何らの誤差なく受信部に伝達される。しかし、オーディオ信号に対する伝送線路が不安定ならば固定ビット率で構成されたデータは受信端で確に解析できなくなる。即ち、

み受取ると、該当されるオーディオデータを復元しにくくなるので再生時相当の音質の劣化を招く恐れがある。

【0006】一般に、デジタルオーディオ符号化方法は1個のビットストリームのヘッダーに1個のビット率に対する情報のみを含んでいる。例えば、1つのビットストリームのヘッダー情報にビット率が128kbpsなら128kbpsのビットストリームを使用し続ける。このような方法の長所は該当ビット率で示せる最上の音質を表現しうることである。即ち、特定ビット率に対し、例えば64kbps、48kbps、32kbpsなどオーディオデータに対する最適のビット文字列を構成する。このような方法の短所は、伝送線路の状態に非常に敏感に作用するということである。伝送線路が非常に安定的なら問題ないが、伝送線路が非常に不安定な場合は正しいデータを再生できない。例えば、1つのオーディオフレームがn個のスロットで構成されていると仮定する場合、与えられた時間内にn個のスロットが全て受信部に伝達されると正しいデータを再生する。しかし伝送線路が不安定なため、与えられた時間内にn-m個のスロットのみ伝達されるとしたら正しいデータが再生できない。また、図3に基づき一つの送信部から提供されるデータを複数の受信部から受取る場合を考えて見よう。各受信部毎に伝送線路の容量が異なったり各受信部により相異なるビット率が要求される時、固定ビット率のみを支援する送信部ならこのような要求事項を満たしにくい。この場合、オーディオのビットストリームが様々な段階のビット率で構成されていると与えられた環境や使用者の要求事項に対して適切に対処し

うる。

【0007】このような目的のためにビット率の調節を可能にする方法は3つに大別される。第1の方法はビットストリームに多段階(Layer)の情報が順次に構成されていて要求するビット率に単にビットストリームのみ切断して伝達する方法である。図4に示されたように基本階層(Base Layer)から最上位階層(Top Layer)までのビットストリームが順次に構成されている。各階層による付加情報(side information)とオーディオデータが全て一つのビットストリームに記録される方式である。従つて、使用者の要求が基本階層のみを要求すると基本階層に該当するビットストリームを伝達し、第1階層(Layer 1)に対する情報を要求すると第1階層までのビットストリームを伝達する。また、最上位階層の情報を要求すればビットストリーム全体を伝達する。第2の方法は送信端と受信端との間に何れの装置(例えば、変換器)が使用者の要求に応じてビットストリームを再構成(reformatting)することである。図5に示されたように符号化装置においては1つのビット率でビットストリームを作成するが、変換器において使用者の要求に応じてさらに低い

作れるようにするための何れの付加情報が必ず存在すべきである。第3の方法は図6に示されたように、変換器において再符号化(reencoding)することである。再符号化とは復号化器(decoder)で行うように完全に信号を復元した後、符号化装置で行うように使用者の要求するビット率で再び符号化して伝達することである。復号化の全ての過程を経てPCMデータを作り、符号化の全ての過程を経て所望のビット率に対したビットストリームを生成して伝達することである。例えば、何れの主伝送線路に64kbpsビットストリームが伝達されるが、何れの使用者に向かう伝送線路の容量が32kbpsとする時、中間の変換器ではまず64kbpsに対する復号化器を用いてPCMデータを作つてから32kbps符号化装置を動作させて32kbpsビットストリームを作つた後、その伝送線路を通してデータを送る。

【0008】前述した方法の中、最初の方法が最適であるが、各階層においてデータの冗長がありうるので性能が劣化する短所がある。第2の方法は最初の方法に比べて音質が多少向上されるが、さらに低いビットストリームの作成は符号化装置から伝達された付加情報に応じて変わる。例えば48kbpsに変換することを所望する場合、元のビットストリームに48kbpsに対する情報がなければ再作成し難くなる。また最初の方法に比べて変換器という段階を経なければならないので時間遅延及び費用が高くつく短所がある。第3の方法は変換器の役割が復号化器と符号化器との役割を同時に行うべきなので複雑度が高くてコスト高及び再符号化に伴う時間遅延(delay)が発生しうる。しかし、変換器の入力に使用されるビットストリームに冗長される部分がないので音質が最初の方法に比べて良い長所がある。前記第2の方法(reformating)と第3の方法(reencoding)との区分は相当曖昧であるが、低ビットストリームを構成する過程に逆量子化(dequantization)過程があれば再符号化と見なす。ビット率の調節可能なシステムにおいて変換器の役割は使用者と送信端との間の単純な連結なので複雑度の面で低いほど良い。従つて、一般に時間遅延のなく、安価の低複雑度の変換器を使用するので制限条件(restriction)として再符号化しない方法を採択する。

【0009】最初の方法のようにビットストリームを作成するためには一般に、図7に示されたように、まず低い階層に対する符号化を行なつた後、再び復号化して得た信号と元の信号との差を再び次の階層の符号化装置の入力として使用して処理する。このような方法は一般に2つ以上の符号化方式を使用する。基本階層を生成するコアコデック(core codec)と次の階層を生成する別のコデック(codec)との混用である。このような方法は2つ以上の符号化装置があるので符号化装置の複雑度を高

は各階層から生成される時間領域のデータを合算してこそ該当階層の正しい時間領域のデータが得られるからである。

【0010】

【発明が解決しようとする課題】本発明が解決しようとする技術的課題は効率的な伝送線路の使用と多様な使用者の要求を満たし、使用者の要求に応じて様々な階層のビット率に対した情報を冗長なしに一つのビットストリームに結合させることにより良い音質のビットストリームを提供するため、ビット率の調節可能なオーディオ符号化/復号化方法及び装置を提供することにある。

【0011】

【課題を解決するための手段】前記技術的課題を解決するための本発明によるビット率の調節可能なオーディオデータ符号化方法は、オーディオ信号を基本階層と所定数の向上階層よりなる階層的なデータストリームに符号化する方法において、入力オーディオ信号を信号処理して所定の符号化帯域別に量子化する量子化処理段階と、前記基本階層に相応する量子化されたデータを所定の階層サイズ内で符号化する基本階層符号化段階と、前記符号化された階層の次の向上階層に属する量子化されたデータ及び前記符号化された階層に属しながら符号化されなく残っている量子化されたデータを所定の階層サイズ内で符号化する階層符号化段階と、前記階層符号化段階を全ての階層に対して行う順次符号化段階とを含むことを特徴とする。

【0012】前記基本階層符号化段階、階層符号化段階及び順次符号化段階の符号化は符号化しようとする該当階層に属する各帯域に対し、前記帯域に相応する量子化されたデータのビット数を示す全域ビット割当情報を求める第1段階と、前記階層の各帯域の大きさの内で各帯域別に割当てられるビット数を求める第2段階と、前記割当てられたビット数に相応する量子化されたデータに対して前記帯域を構成している所定周波数成分別に量子化されたデータの存在有無を示す位置情報を生成する第3段階と、前記全域ビット割当情報、量子化ステップサイズ、位置情報及び前記帯域別に割当てられたビット数に相応する量子化されたデータを所定の符号化方法によりビットストリームに生成する第4段階となる。前記帯域別に割当てられたビット数に相応する量子化されたデータは、量子化されたデータの最上位ビットから下位ビット順に相応する量子化されたデータであることが望ましい。前記第4段階の前記帯域別に割当てられたビット数に相応する量子化されたデータの符号化は、前記量子化されたデータを所定数のビット単位に束ねて符号化することが望ましい。前記第2段階の割当てられたビット数は1ビットであり、前記第4段階は前記全域ビット割

号化であることが望ましい。そして、前記量子化処理段階は時間領域の入力オーディオ信号を周波数領域の信号に変換する段階と、前記時間/周波数マッピングにより周波数成分に変換された入力オーディオ信号を所定の帯域信号に束ねて、前記各帯域でのマスキングスレショルドを計算する段階と、各帯域の量子化ノイズがマスキングスレショルドより小さくなるように所定の符号化帯域別に量子化する段階よりなることが望ましい。

【0013】前記技術的課題を解決するための本発明によるビット率の調節可能なオーディオデータ符号化装置は、オーディオ信号を所定数の異なるビット率よりなる階層的なビット率を有するように符号化する装置において、入力オーディオ信号を信号処理して所定の符号化帯域別に量子化する量子化処理部と、前記基本階層に属する各帯域の量子化されたデータのビット数を示す基本階層の帯域別全域ビット割当情報及び前記各帯域別に割当てられるビット数に相応する量子化されたデータの所定の周波数成分別データの存在有無を示す位置情報、量子化ステップサイズ及び量子化されたデータを符号化し、前記基本階層に対する符号化が終わると、その次の階層に対する全域ビット割当情報、位置情報、量子化ステップサイズ及び量子化されたデータを符号化し、これを全ての階層に対して行ってビットストリームを形成するビットパッキング部とを含むことが望ましい。前記帯域別に割当てられたビット数に相応する量子化されたデータの符号化は、量子化されたデータの最上位ビットから下位ビット順に相応する量子化されたデータを符号化することが望ましい。

【0014】前記ビットパッキング部の帯域別に割当てられたビット数に相応する量子化されたデータの符号化は、前記量子化されたデータを所定数のビット単位に束ねて符号化することが望ましい。前記ビットパッキング部の帯域別に割当てられるビット数は1ビットであり、前記符号化は全域ビット割当情報と位置情報を所定の符号化方法により符号化することが望ましい。前記ビットパッキング部の符号化は低周波数成分から高周波数成分順に行なうことが望ましい。前記量子化処理部は時間領域の入力オーディオ信号を周波数領域の信号に変換する時間/周波数マッピング部と、前記時間/周波数マッピングにより周波数成分に変換された入力オーディオ信号を所定の帯域信号に結合し、前記各帯域でのマスキングスレショルドを計算する聴覚心理部と、各帯域の量子化ノイズがマスキングスレショルドより小さくなるように所定の符号化帯域別に量子化する量子化部とを含むことが望ましい。

【0015】前記技術的課題を解決するための本発明によるビット率の調節可能なオーディオデータ復号化方法

示す全域ビット割当情報、量子化ステップサイズ、前記割当てられたビット数に相応する量子化されたデータの周波数成分別存在有無を示す位置情報及び量子化されたデータを前記ビットストリームの階層が生成された順序通り復号化するが、前記各階層に属する帯域に割当てられたビット数を求めて前記割当てられたビット数に相応する量子化されたデータを前記位置情報に参照して復号化する復号化段階と、前記復号化された量子化ステップサイズと量子化されたデータを元の大きさの信号に復元する段階と、前記逆量子化された周波数領域のオーディオ信号を時間領域の信号に変換する段階とを含むことが望ましい。前記復号化段階の量子化されたデータの復号化は量子化されたデータの最上位ビットから下位ビット順に符号化されたデータを復号化することが望ましい。前記復号化段階の量子化されたデータの復号化は所定数のビットよりなるペクトル単位で符号化されたデータを復号化することが望ましい。

【0016】前記技術的課題を解決するための本発明によるビット率の調節可能なオーディオデータ復号化装置は、階層的なビット率を有するように符号化されたオーディオデータを復号化する装置において、前記階層に属する各帯域に相応する量子化されたデータのビット数を示す全域ビット割当情報、量子化ステップサイズ、前記割当てられたビット数に相応する量子化されたデータの周波数成分別存在有無を示す位置情報及び量子化されたデータを前記ビットストリームの階層が生成された順序通り復号化するが、前記各階層に属する帯域に割当てられたビット数を求めて前記割当てられたビット数に相応する量子化されたデータを前記位置情報に参照して復号化するビットストリーム分析部と、前記ビットストリーム分析部から復号化された量子化ステップサイズと量子化されたデータを元の大きさの信号に復元する逆量子化部と、前記逆量子化部から逆量子化された周波数領域のオーディオ信号を時間領域の信号に変換する周波数/時間マッピング部とを含むことが望ましい。

【0017】

【発明の実施の形態】以下、添付された図面に基づき本発明の望ましい一実施の形態を詳しく説明する。本発明は様々な階層のビット率に対する情報を一つのビットストリームで表現する方法であって、使用者の要求または伝送線路の状態に応じて各段階のビット率に対したビットストリームを簡単に再構成して送れる。例えば、基本階層は16kbps、最上位階層は64kbps、各階層が8kbpsで構成されていると仮定する。そうすると符号化装置で構成されるビットストリームは最上位階層の64kbpsに対するビットストリーム内に各階層16、24、32、40、48、56、64kbpsに対する情報が共に貯蔵されている形態であ

するデータを要求すればただ前部のビットストリームのみを切取って伝達する。

【0018】図8は本発明によるビット率の調節可能なオーディオ符号化装置の望ましい一実施の形態に対する構成を示したブロック図であって、量子化処理部830及びビットパッキング部840よりなる。前記量子化処理部830は入力オーディオ信号を信号処理して所定の符号化帯域別に量子化するブロックであって、時間/周波数マッピング部800、聴覚心理部810及び量子化部820よりなる。前記時間/周波数マッピング部800は時間領域の入力オーディオ信号を周波数領域の信号に変換する。時間上で人間の認知する信号の特性差があまり大きくないが、このように変換された周波数領域の信号は人間の聴覚心理モデルに応じて各帯域で人間が感じられる信号と感じられない信号との差が大きいため各周波数帯域に応じる量子化ビットを異に割当てることにより圧縮の効率を高めうる。

【0019】聴覚心理部810は前記時間/周波数マッピング部800により周波数成分に変換された入力オーディオ信号を所定の帯域信号に束ね、各信号の相互作用により発生されるマスキング現象を用いて各帯域におけるマスキングスレショルドを計算する。量子化部820は各帯域の量子化ノイズがマスキングスレショルドより小さくなるように所定の符号化帯域別に量子化する。即ち、人間が聞いても感じられないように各帯域の量子化ノイズの大きさが前記マスキングスレショルドより小さくなるように各帯域の周波数信号をスカラー量子化する。聴覚心理部(810)で計算したマスキングスレショルドと各帯域から発生するノイズの比率であるNMR(Noise-to-Mask Ratio)を用いて全域のNMR値が0dB以下になるように量子化する。NMR値が0dB以下ということは量子化ノイズに比べてマスキング値が高いことを示し、これは量子化ノイズを人間が聞取れないということを意味する。

【0020】前記ビットパッキング部840は量子化部820から量子化された周波数帯域のデータ値を各階層に該当するビット率に応じて該当帯域の付加情報とオーディオデータに対する量子化情報を結合してビットストリームを生成する。これをさらに詳しく説明すれば、符号化しようとする該当階層に属する各帯域に対し、前記帯域に相応する量子化されたデータのビット数を示す全域ビット割当情報を求める。それから前記階層の各帯域の大きさ内で各帯域別に割当てられるビット数を求め、前記割当てられたビット数に相応する量子化されたデータに対して前記帯域を構成している所定の周波数成分別に量子化されたデータの存在有無を示す位置情報を生成する。次いで、前記全域ビット割当情報、量子化ステップサイズ、位置情報及び前記帯域別に割当てられたビット数に

は量子化されたデータの最上位ビット(MSB)から下位ビット順に相応する量子化されたデータを符号化する。そして、前記量子化されたデータの符号化は所定数のビット単位に束ねて低周波数成分から高周波数成分順に符号化することもできる。

【0021】一方、本発明の動作を説明すると次の通りである。1つのビットストリーム内に多層の情報を共に記録する最も一般的な方法は図7に示されたように、低い階層に符号化した結果を貯蔵し、この結果を復号化して再生された信号と元の信号との差を次の階層で処理して貯蔵する方式である。例えば、基本階層を16kbpsとしよう。それでは、まず16kbpsに符号化した結果をビットストリームに貯蔵する。そして、元の信号と16kbpsに符号化されたデータを復号化した結果の差を元の信号と仮定して次の階層の符号化を行なう。このような方式はビットストリームに多くの冗長(redundancy)があるために音質が劣る短所がある。即ち、低周波数帯域のデータは多層で冗長されて表現される。これ为了避免するためには同じ周波数成分に対して各階層で表現されるデータの間に連続性を与えるべきである。例えば、何れの周波数成分のデータに対してi番目の階層のビット割当がmビットでi-1番目の階層のビット割当がnビットなら、(n+m)ビットに表現されたことが1つの情報として効果があるべきである。

【0022】本発明では同じ周波数成分に対して各階層で表現されるデータの間の連続性のためにデータのMSBから記録する方式を使用する。図9に基づいて例を挙げて説明する。量子化部820でNMR0に対して量子化した結果が図9(A)としよう。データは各階層でMSBから表現される。即ち、基本階層では図9(B)に示されたように最上位MSBのみが表現され、次の階層はその次のMSBが表現される。このような方法で最上位階層まで進行されるとデータの全てのビットが表現される。従って、階層が進行されるほどさらに細密な情報を示す。復号化部に伝達されるデータはこの帯域に割当てられるビットの長さを示す全域ビット割当情報とデータの位置情報(index)そしてデータビットである。前記位置情報はデータビットが帯域内の何れの位置に存在するかを示す役割をする。これはビットプレーンマスク(bitplane mask)を用いて生成する。ビットプレーンマスクは全てのビットが1の値である。前記位置情報は図10に示されたように前記マスクとデータのMSBをマスクのビット数だけANDに連結して生成される値である。

【0023】前記図10の例ではこの帯域のビット割当情報は7ビットである。最上端のデータを例とすると、基本階層においてデータのMSB1ビットのみあっても復号化器でこの値は'1000000'と認識する。これはこの帯域

値は‘1100000’と認識される。位置情報はデータの位置を表現するだけでなく1というデータ値を示す。従って、図9において‘1000001’を表現する時は位置情報による値1をデータ値から引いた後に使用する。従って、位置情報の1とデータビットにより表現された‘1000000’を足した‘1000001’が復号化器で再生される。図9において2段階と5段階以後のデータは復号化器に伝達しなくとも全てのデータを示しうる。位置情報の表現はハフマン符号を使ったりランレンジス符号化方式または算出符号化方式を使用する。データは該当されるビットをそのまま使用する。またデータビットが1ビットの場合、位置情報のみでデータビットまで示しうる。これはマスクにより生成される1という値は該当位置にMSBが存在するということを意味する。従って追加のデータビットがなくても全体値を表現しうる。例えば、ビット割当

情報が5の場合、基本段階の位置情報にある値1は1と‘10000’を同時に示すので復元される値は‘10001’となる。前述したように1つの階層で1ビットずつ表現することは最も基本的な方式である。これを変形して1つの階層で1ビット以上を表現しうる。各帯域のデータを表現する方法はまず各階層に該当する帯域のMSBを最小限1ビットずつ割当てて表現し、ビット割当情報の大きさ順に低域から高域まで順次に表現する。即ち、ビット割当情報の大きな帯域は重要な帯域なので多くの部分を先に表現してビット割当情報の少ない帯域は相対的に重要度が低いので後で表現する。次のアルゴリズムは各階層のデータ生成アルゴリズムであって、スケーラブルビットストリーム (scalable bitstream) を構成するために各階層のビット割当情報とデータを計算する部分である。

【0024】

```

for(band=0;band<;band_end[Top_Layer];band++)
    for(layer=Base_Layer;layer<;Top_Layer;layer++)
        layer_bal[layer][band]=0;
/*max_bit_alloc:量子化部で計算された該当帯域のビットの大きさ*/
        for(band=0;band<;band[Top_Layer];band++)
            bit_alloc[band]=max_bit_alloc[band];
        used_bits=0;
        for(layer=Base_Layer;layer<;Top_layer;layer++)
        {
            for(band=band_start[layer];band<;band_end[layer];band++)
            {
                if(bit_alloc[band]<;2 continue;
                layer_bal[layer][band]=bit_alloc[band]/2;
                bits=quantizer(isample,
                    layer_sample[layer],
                    bit_alloc[band],
                    layer_bal[layer][band],
                    band);
                if(used_bits+bits>;available_bits_for_this_layer[layer])break;
                used_bits+=bits;
                bit_alloc[band]-=layer_bal[layer][band];
            }
            while(used_bits<;available_bits_for_this_layer[layer])
            {
                band=-1;
                max_bal=0;
                for(scfb=0;scfb<;band_end[layer];scfb++)
                {
                    if(max_bal<;bit_alloc[scfb])
                    {
                        max_bal=bit_alloc[scfb];
                        band=scfb;
                    }
                }
            }
        }
    }
}

```

```

        bit_alloc[band]--;
        bits=quantizer(isample,
        layer_sample[layer],
        max_bit_alloc[band],
        layer_bal[layer][band],
        band);
        if(used_bits+bits>available_bits_for_this_layer[layer])
        {
            layer_bal[layer][band]--;
            bit_alloc[band]++;
            break;
        }
        used_bits+=bits;
    }
    for(band=0;band<band_end[layer];band++)
        max_bit_alloc[band]-=layer_bal[layer][band];
}

```

【0025】図11及び図12は前記アルゴリズムを行った一例の中、基本階層と階層1に対した処理結果を示す。図11及び図12において初期ビット割当情報は全域ビット割当情報の1/2に該当する情報である。その後、残っているビット割当情報（全域ビット割当情報－初期ビット割当情報）のうち大きさ順にビット割当を実施する。前記アルゴリズムの例では基本階層で許容可能なビット内に2番目の帯域までの追加ビットが与えられることを示す。基本階層の処理が終わった後、元の信号から処理された信号を引いた信号を次の階層で処理する。このように階層が進行されるに応じて細密なデータ値の表現が行われるので階層が上がることにより良質の音質を示しうる。このように表現されたデータを用いてスケーラブルビットストリームを構成する方法は次の通りである。まず基本階層の帯域幅に対した付加情報を記録する。付加情報は各帯域のビット割当情報と逆量子化のために必要な量子化ステップサイズである。ここで、量子化ステップサイズは該当帯域に初めてビット割当られる場合に記録する。付加情報の後に基本階層に対するオーディオデータ値を記録する。データ値の記録は図11、図12のアルゴリズムにより最初帯域から基本階層のビット率に応じる許容可能なビット数内で表現可能な帯域までのMSB情報とデータの位置情報を計算してから順次に記録する。図11、図12の例において、ビットストリーム内に記録されるのは位置情報と実線で四角に表示されたデータビットである。この際、各階層のビット割当情報が0であるのはどんなデータも記録しない。また、各階層のビット割当情報が0でない場合、位置情報に何等の情報もなければ(即ち、位置情報が全て0なら)データビットは記録しなく位置情報のみ記録す

3は記録されたビットストリームの構造である。

【0026】一方、本発明による符号化方法をさらに具体的な例に基づき詳しく説明すれば次の通りである。基本階層は16kbps、最上位階層は64kbps、各階層は8kbpsで構成される場合を例とする。即ち、16、24、32、40、48、56、64kbpsの7階層のスケーラブルビットストリームを構成する例である。表1に示されたように、各階層は処理される帯域幅が制限されている。入力データは48kHzでサンプリングされたPCMデータであり、1フレームの大きさは1024個である。64kbpsビット率の1フレームで使えるビットの数は平均1356ビット(64000 bits/sec*(1024/48000sec))である。

【0027】

【表1】

ビット率(kbps)	帯域幅(khz)
16	3.5
24	6
32	7.5
40	10
48	12
56	14
64	16

まず入力データは聴覚心理モデルを通して必要な情報を計算する。聴覚心理モデルを計算する方法はISO/IEC 11172-3のモデル2を使用する。聴覚心理部810の出力は現在処理されているフレームのブロックタイプ(ロング、スタート、ショート、ストップ)と表2に示された各処理

る。

【0028】

【表2】

帯域	周波数(Hz)	帯域	周波数(Hz)	帯域	周波数(Hz)
0	188	10	3093	20	9843
1	375	11	3468	21	10593
2	562	12	3843	22	11343
3	844	13	4593	23	12093
4	1125	14	5343	24	13031
5	1406	15	6093	25	14062
6	1687	16	6843	26	15093
7	1968	17	7593	27	16125
8	2343	18	8343	28	
9	2718	19	9093	29	

聴覚心理部810の出力のブロックタイプによりT/Fマッピング部800ではMDCT(Modified Discrete Cosine Transform)を用いて時間領域のデータを周波数領域のデータに変換する。この際ブロックの大きさは、ロングブロックの場合2048であり、ショートブロックの場合256であるMDCTを8回行う。周波数領域に変換されたデータは聴覚心理部810の出力値のSMRを用いて量子化される。量子化は非線形スカラー量子化(non-linear scalar quantization)を使用し、基本的な量子化ステップサイズは21/4を使用する。量子化はNMR値が0以下となるように行う。この際、得られる出力は量子化されたデータと各処理帯域との量子化ステップサイズ、そしてこの帯域の最大値を表現しうるビット数(全域ビット割当情報)である。

【0029】各階層のデータは前記アルゴリズムを用いて生成する。これを実施の形態に基づき説明すれば次の通りである。最下位階層の帯域幅は3.5kHzなので処理帯域は12番目の帯域までである。まず12番目の帯域までの全域ビット割当情報を用いて基本階層のビット割当情報を生成する。0番目の帯域から全域ビット割当情報の半分のビット内にあるデータを読み出でして基本階層の許容可能なビット内に表現できるかを調べる。これが可能ならその次の帯域のデータに対して0番目の帯域の方法を用いて調べる。このように12番目の帯域まで進行するうち許容可能なビット数を越える帯域があったり、12番目の帯域まで全て許容可能なビット数内で表現できれば次は現在処理されている帯域のうちビット割当情報の大きい順に1ビットずつ割当てて調べる。位置情報値はハフマン符号とランレンジス符号のうち小さいのを使用する。基本階層の処理が終わると、基本階層で処理されたデータ値を元のデータ値から引いた値を次の階層の入力データとして使用する。他の階層の処理も基本段階と同様に

る。図14は前記復号化装置の構成を示したブロック図であって、ビットストリーム分析部10、逆量子化部20及び周波数/時間マッピング部30よりなる。前記ビットストリーム分析部10は前記階層の全域ビット割当情報、量子化ステップサイズ、位置情報及び量子化されたデータを前記ビットストリーム階層の生成順に応じて復号化するが、前記各階層に属する帯域に割当てられたビット数を求める、前記割当てられたビット数に相応する量子化されたデータを前記位置情報に基づき復号化する。前記逆量子化部20は復号化された量子化ステップサイズと量子化されたデータを用いて元の大きさの信号に復元する。前記周波数/時間マッピング部30は周波数領域のオーディオ信号を再び時間領域の信号に変換して使用者が再生できるようにする。

【0031】次いで、前記復号化装置の動作を説明する。前述した符号化装置により生成されたビットストリームの復号化過程は符号化過程の逆順である。まず基本階層に対する情報を復号化する。簡単にその過程を説明すれば、まず基本階層の付加情報、即ち全域ビット割当情報、量子化ステップサイズ及び位置情報を復号化する。そして前記基本階層に属する帯域に割当てられたビット数を求める、前記割当てられたビット数に相応する量子化されたデータを前記位置情報に基づき復号化する。ここで前記量子化されたデータの最上位ビットから下位ビット順に、そして低周波数から高周波数順に符号化されたビットストリームで量子化された値を復号化する。前記復号化は符号化段階で使用した符号化方法に相応する復号化方法を使用する。望ましくは無損失復号化方法が使われ、その例としてハフマン復号化方法及び算術復号化方法が使われる。このように基本階層に割当てられた大きさのビットストリームに対する復号化が終わる

されたデータは前記符号化器と逆順に、図14に示された逆量子化部20と周波数/時間マッピング部30を経て復元された信号を作れる。

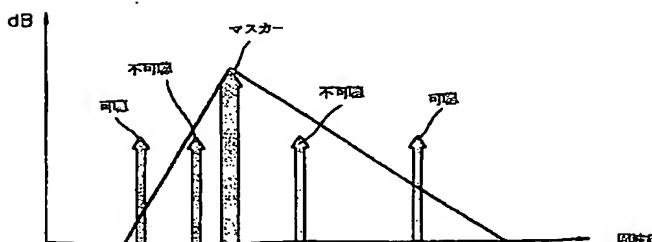
【0032】

【発明の効果】本発明によれば、多層のビット率に対しデータを一つのビットストリーム内に表現することにより、伝送線路の状態または使用者の要求事項に対して適応的に対処しうる。また、ビット率の調節可能な形態のビットストリームを生成するので一つのビットストリームで多層のビットストリームを生成しうる。そして各階層に全て同一な符号化器を使用するので符号化器の複雑度が低い。また各階層で時間領域のデータ差を符号化せず周波数領域のデータを処理するので符号化器の複雑度が低い。そして、各階層に全て同一な復号化器を使用するので復号化器の複雑度が低い。また、各階層に時間領域のデータをただ一回のみ生成するので復号化器の複雑度が低い。そして、多層のビットストリームを簡単に生成しうるので変換器の複雑度が低い。また、各階層のデータビットの間にビット連続性を有するので良い音質が提供でき、伝送線路の状態及び使用者の要求に容易に対応しうる。そして、本発明は多様な使用者の要求に相応するためにビットストリームを柔軟に構成する。即ち、使用者の要求に応じて多層のビット率に対した情報を冗長なしに一つのビットストリームに結合させることにより良い音質のビットストリームを提供しうる。また、送信端と受信端との間に何等の変換器も必要なく伝送線路の状態や使用者のいかなる要求も受け入れられる長所がある。そして、本発明の適用分野は有線または無線を使用して提供される全ての応用分野である。

【図面の簡単な説明】

【図1】聴覚心理特性を用いた一般のオーディオデータ符号化装置の構成を示したブロック図である。

【図2】



【図2】マスキング現象を説明するための図である。

【図3】一般の伝送線路の構成を示したブロック図である。

【図4】多段階のビット率に対した情報を含んでいるビットストリームを示した図である。

【図5】再作成を説明するための図である。

【図6】再符号化を説明するための図である。

【図7】ビット率の調節可能な符号化器の一般の構造を示したブロック図である。

【図8】本発明による符号化装置の構成を示したブロック図である。

【図9】(A)から(G)は、本発明による符号化装置の動作原理を説明するための一例を示した図である。

【図10】ビットプレーンマスクを用いた位置情報抽出を説明するための図である。

【図11】各階層のデータ生成アルゴリズムの実行例を示した図である。

【図12】各階層のデータ生成アルゴリズムの実行例を示した図である。

【図13】本発明により作成されたビットストリームの構造を示した図である。

【図14】本発明による復号化装置の構成を示したブロック図である。

【符号の説明】

10 ビットストリーム分析部

20 逆量子化部

30 周波数／時間マッピング部

800 時間／周波数マッピング部

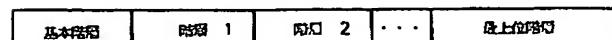
810 聴覚心理部

820 量子化部

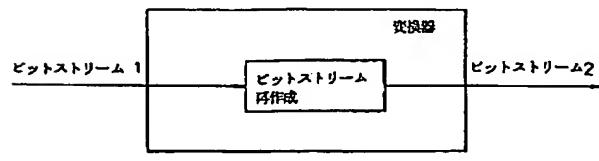
830 量子化処理部

840 ビットパッキング部

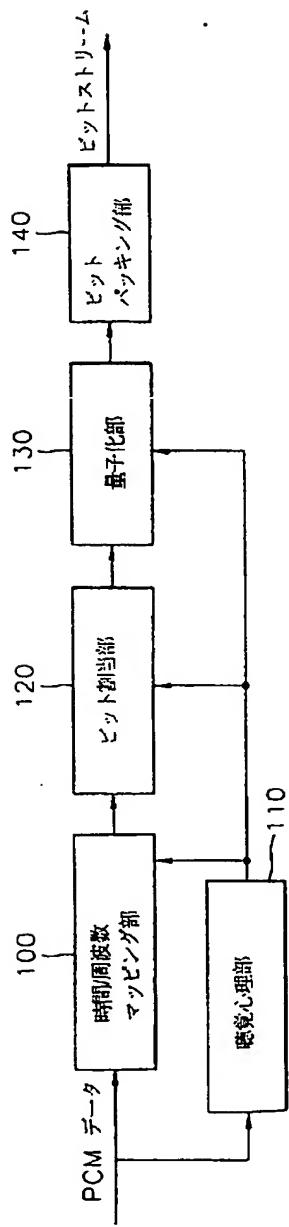
【図4】



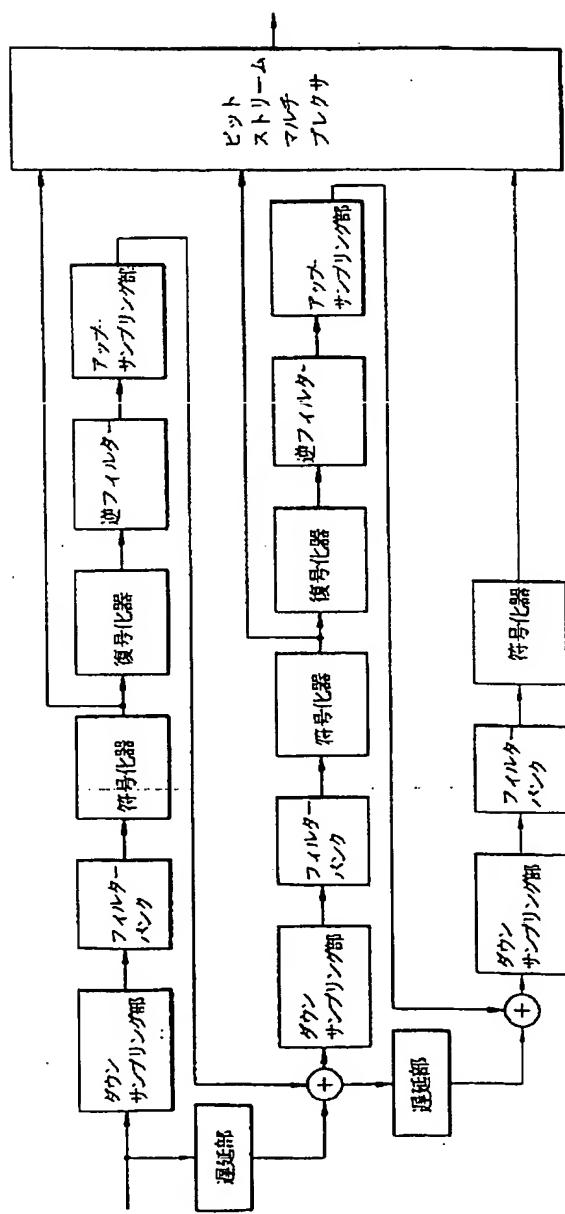
【図5】



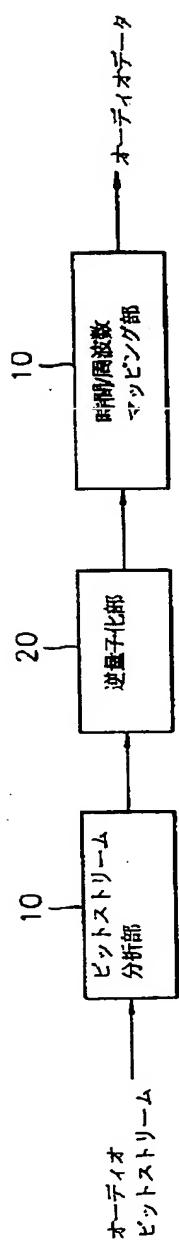
【図 1】



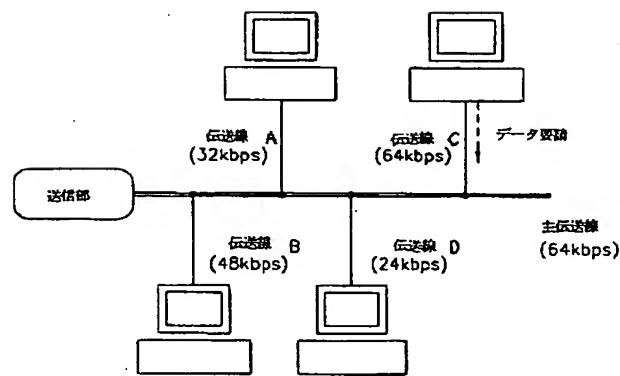
【図 7】



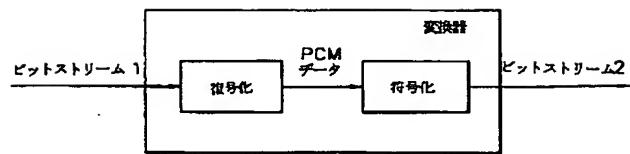
【図 1-4】



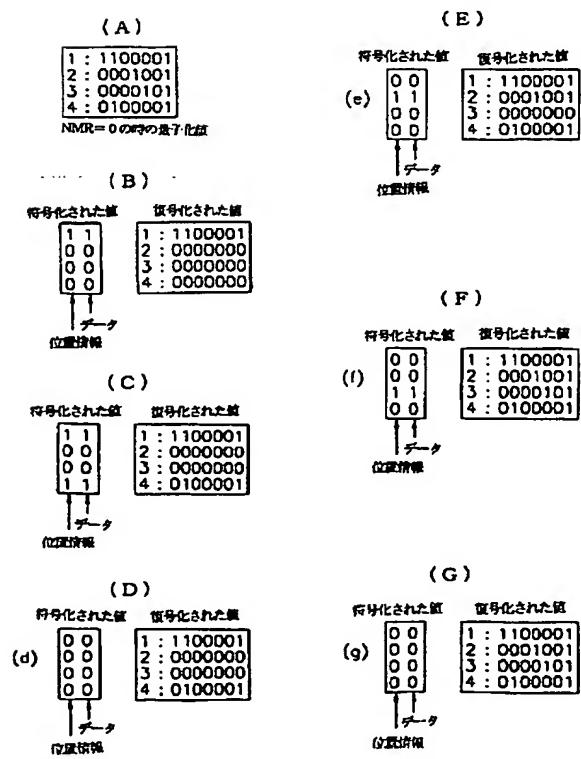
【図3】



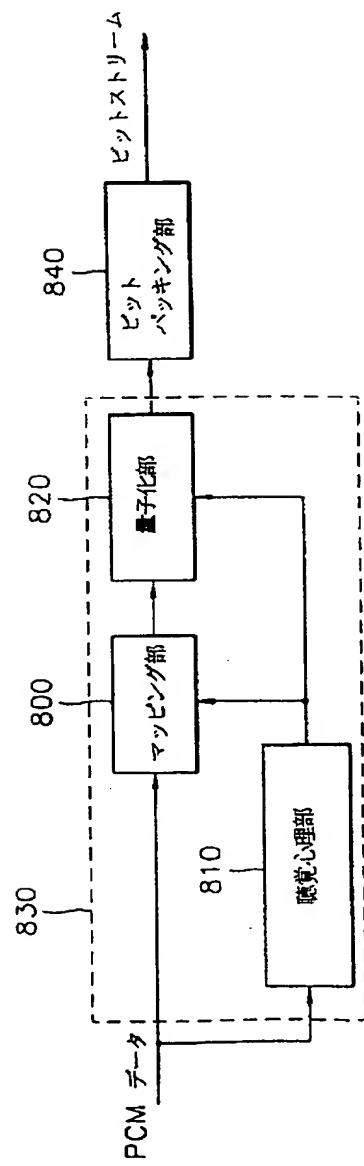
【図6】



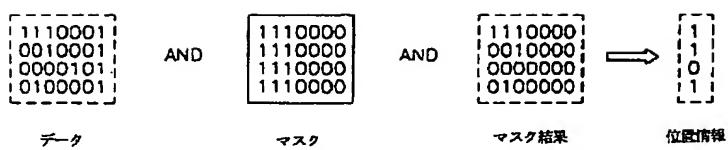
【図9】



【図 8】



【図 10】



【図 11】

領域番号	全域ビット 割当情報	初期データ	初期ビット 割当情報	位置情報	データ ビット	基本対応ビット 割当情報	位置情報	データビット
1	7	1000001 1000001 1000001 100	3	1 1 1 0	1000000 1000000 1000000 100	4	1 1 1 0	1000000 1000000 1000000 11
2	6	1000001 1000001 1000001 101	3	1 1 1 0	1000000 1000000 1000000 101	4	1 1 1 1	1000000 1000000 11111 100
3	5	1000001 1000001 1000001 101	2	1 1 0 0	1000000 1000000 1000000 101	2	1 1 0 0	1000000 1000000 11111 101
4	4	111 101 1100 1	2	0 1 1 0	111 100 1011 1	2	0 1 1 0	111 100 1011 1

【図 12】

全域ビット 割当情報	初期データ	初期ビット 割当情報	位置情報	データ ビット	1回目ビット 割当情報	位置情報	データビット
1	3	000 000 000 0	0 0 0 0	0 0 0 0	1 0 0 0	0 0 0 0	0 0 0 0
2	2	100 000 000 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
3	3	110 000 000 0	0 0 0 0	0 0 0 0	1 0 0 0	0 0 0 0	1 1 1 0
4	2	110 000 000 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
5	5	10001 101 000 0	2 0 0 0	10000 101 000 0	3 0 0 0	0 0 0 0	10000 100 000 0
6	4	1001 1000 1000 1	2 0 0 0	1000 1000 1000 1	2 0 0 0	0 0 0 0	1000 1000 1000 1
7	4	1001 1000 1000 1	2 0 0 1	1001 1000 1000 1	2 0 0 1	0 0 0 1	1001 1000 1000 1

【図13】

